Dr Milica Šutova, vanredni profesor
Pravnog fakulteta Univerziteta Goce Delčev u Štipu
Republika Severna Makedonija
ORCID: 0009-0009-2972-1091

Dr Ksenija Paunović, naučni saradnik
Univerziteta u Kragujevcu
ORCID: 0000-0002-2064-9419

Izvorni naučni rad
UDK: 004.8:34(4-672EU)
DOI: https://doi.org/10.46793/XXIIMajsko.573S

DRUŠTVENI RIZICI PRIMENE VEŠTAČKE INTELIGENCIJE I NJENO REGULISANJE U PRAVU EVROPSKE UNIJE

Rezime

Veštačka inteligencija se primenjuje u gotovo svim oblastima života. Pravna regulativa ne može da prati njen brz i razgranat razvoj. Dosadašnje iskustvo u vezi sa primenom postojeće slabe inteligencije dovodi do uočljivih i ozbiljnih društvenih rizika. Iz tog razloga u Evropskoj uniji donet je Zakon o veštačkoj inteligenciji kojim je regulisan načelni odnos društva prema ovom fenomenu. Zakon je zasnovan na ozbiljnosti rizika u primeni, pa određene oblike smatra zabranjenim, a dozvoljava veštačku inteligenciju visokog rizika, postavljajući visoke uslove za njeno prihvatanje. Pri tome, predmet ovog zakona nije definisanje prava, obaveza ili pravne zaštite koja bi se odnosila na samu veštačku inteligenciju kao subjekta prava

Ključne reči: *slaba veštačka inteligencija, jaka veštačka inteligencija, zabranjena veštačka inteligencija, sistemi visokog rizika, transparentnost.*

1. Uvod

Veštačka inteligencija utiče danas na gotovo sve oblasti čovekovog života. Svaka primena, npr. u kulturi, politici, vojnoj industriji, klasičnom autorskom pravu i patentnom pravu izaziva niz pitanja i postavlja nove izazove.

Primena veštačke inteligencije, u takvoj meri, nosi sa sobom i određene opasnosti po društvo i čoveka. One moraju biti posebno analizirane u cilju svođenja na društveno prihvatljivu meru, koja se pre svega postiže, adekvatnim pravnim oblikovanjem. Međutim, pravne norme su sistem koji po svojoj prirodi ne može da

prati brzi i eksplozivni tehnološki razvoj. Ta inercija se, pre svega, ogleda u aktivnostima zakonodavstava koja pokazuju oprezan i gotovo statičan prilaz novonastalim problemima. Njihovo rešavanje je ostavljeno sudskoj praksi i njenoj fleksibilnosti, a izlaz se često traži u prilagođavanju već postojećih rešenja u drugim oblastima. Granice sudske prakse su ipak ograničene, jer je njen predmet samo tumačenje prava, a ne i kreiranje, koje isključivo pripada zakonodavcu. Bogata i razvijena sudska praksa, naročito najviših sudova, stvara osnovu za zakonodavnu intervenciju koja mora uzeti u obzir i očekivani tehnološki razvoj.

Zakonodavstvo Evropske unije učinilo je prvi, veoma značajan i ozbiljan korak na pravnom regulisanju veštačke inteligencije. Zakonom o veštačkoj inteligenciji Evropske unije rešeno je osnovno pitanje koje se tiče odnosa društva prema njoj. U tom kontekstu, neki oblici se smatraju neprihvatljivim za društvo, s obzirom na rizik koji se izaziva. Drugi oblici jesu visoko rizični, ali su dozvoljeni. Treći oblici su oni kojima se ne izaziva rizik. Drugim rečima, odnos društva prema veštačkoj inteligenciji određuje se s obzirom na visinu izazvanog rizika. U okviru ovakve podele uzeti su u obzir elementi vezani za širu organizaciju društvenih odnosa, zaštitu ličnosti, privatnosti, zaštitu osnovnih ljudskih prava, naročito u oblasti osetljive materije krivičnog prava i sudskog postupka, ali i pitanje podrške tehnološkom razvoju.

Predmet rada je predstavljanje opasnosti koje nosi veštačka inteligencija. Pored toga, razmotrena su i neka bitna pitanja iz zakona Evropske unije. Reč je o vrlo kompleksnim i komplikovanim pravnim normama, pa su ostavljeni i dopunski rokovi za njihovu primenu. Tako će se odredbe o visoko rizičnoj inteligenciji primenjivati 36 meseci posle stupanja na snagu Zakona o veštačkoj inteligenciji.

2. Pojam i vrste veštačke inteligencije

Ne postoji jedinstvena definicija veštačke inteligencije, jer se pojam upotrebljava usko vezan za kontekst. Ona se može podeliti po različitim kriterijumima u zavisnosti od njene strukture i načina funkcionisanja. Jedna od definicija je: „Veštačka inteligencija opisuje sposobnost mašina bazirana na algoritmima da zadatke izvede autonomno i da pri tome prilagodljivo može reagovati na nepoznate situacije.“[1] Postupanje je slično čovekovom. Ona ne izvršava samo repetitivne zadatke, nego uči iz uspeha i neuspeha i shodno tome prilagođava postupanje.

U budućnosti, veštački inteligentne mašine treba da budu u situaciji da razmišljaju kao ljudi i da na taj način komuniciraju. Međutim, s obzirom na aktuelne kriterijume, ali i moguće buduće opasnosti, ukazaćemo samo u osnovnim crtama na karakteristične podele.[2]

[1] Philip Schurr, mindsquere.de/knowhow, 15.06.2024.

[2] Artenvon KI: Definition, Abgrenzung&Anwendung, Otris software, dostupno na: https://www.otris.de/wiki/arten-von-ki/, 10.4.2025.

2.1. Slaba veštačka inteligencija

Načelno se razlikuju dva različita tipa, a to su slaba i jaka veštačka inteligencija. Slaba veštačka inteligencija je prva forma veštačke inteligencije koja se, u ograničenim područjima, danas bliži ljudskoj inteligenciji. Pri tome se inteligentni sistemi svakako ograničavaju na konkretna područja primene. Tako, na primer, neka veštačka inteligencija može da oblikuje perfektne tekstove, ali ne može da komunicira niti da rešava ukrštene reči ili bilo šta drugo. Slaba veštačka inteligencija nije u stanju da ostvari dublje razumevanje postavljenih problema, pa ostaje na površinskoj ravni i nije u mogućnosti da dosegne dublje razumevanje za rešenje problema koji je dat u zadatak. Ona se služi matematičkim metodama i informatikom koji su specijalno razvijeni za određene zahteve. Slabi sistemi su već prisutni u svakodnevnom životu. Oni se nalaze u pozadini programa kojima se raspoznaju oznake i tekstovi, navigacionih sistema, prepoznavanja jezika, individualnih predstavljanja reklamnih kampanja[3].

2.2. Jaka veštačka inteligencija

Jaka veštačka inteligencija se često označava i kao super inteligencija. To su sistemi koji dostižu čovekove intelektualne sposobnosti, pa čak ih i prevazilaze. Oni samostalno deluju inteligentno i fleksibilno i nisu ograničeni samo na rešenje konkretnog problema. Međutim, ovaj nivo razvoja veštačke inteligencije još uvek nije razvijen i predmet je brojnih diskusija o tome da li je uopšte moguć. Ukoliko se u nekom trenutku i ostvari, ona bi pokazivala sledeća svojstva: logično razmišljanje, sposobnost za odlučivanje, sposobnost planiranja i učenja, komunikacije na prirodnom jeziku, kao i kombinaciju svih sposobnosti da bi se ostvario viši cilj.

Ova osnovna podela veštačke inteligencije pruža samo osnovne pojmove, pa kao takva, ne ulazi u pitanje načina funkcionisanja veštačke inteligencije, njenog treniranja i specifičnih upotreba[4].

2.3. Vrste veštačke inteligencije prema načinu funkcionisanja i sposobnostima

Podela na slabu i jaku inteligenciju je opšta podela, dok se drugom podelom ukazuje na detaljnu klasifikaciju. Ona obuhvata različite oblike veštačke inteligencije, a zasniva se na načinu funkcionisanja i njenoj sposobnosti. Postoje 4 specifične kategorije u okviru ove podele[5].

Prva su reaktivne mašine (*Reactive Machines*) koje su najosnovnija forma veštačke inteligencije, a kategorizuju se kao slaba veštačka inteligencija. Ove forme

[3] Philip Schurr, mindsquere.de/knowhow, 16.06.2024.
[4] Philip Schurr, mindsquere.de/knowhow, 24.01.2024.
[5] *Isto.*

su programirane da reaguju na specifične podatke, a da ne mogu da memorišu prošle događaje ili da ih koriste[6]. Drugim rečima, ovaj oblik veštačke inteligencije ne može da koristi iskustva iz prošlosti za sadašnje odluke. Očigledni primer za to je šah kompjuter *Deep Blue* koji je 1997. godine pobedio svetskog prvaka Gari Kasparova. *Deep Blue* analizira sve moguće šahovske poteze koji se zasnivaju na aktuelnoj situaciji na šahovskoj tabli i bira najbolji potez, a da pri tome ne uzima u obzir ranije igre i ne uči iz njih[7].

Druga kategorija je ograničena mogućnost memorisanja (*Limited Memory*). Ovaj vid ide korak dalje, i može da koristi podatke iz prošlosti u cilju donošenja budućih odluka[8]. Ovi modeli imaju ograničenu mogućnost da memorišu istorijske podatke i da ih koriste. Primer za to su automobili koji sami uče, koji mogu da uzmu u obzir podatke iz saobraćaja i uslove okruženja da bi sigurno upravljali. Ovi automobili znaju kako se drugi normalno ponašaju u saobraćaju, kako izgleda čovek ili biciklista i poznaju saobraćajna pravila. Kada je reč o novoj i do sada nepoznatoj situaciji, ova vrsta je memoriše i zna na koji način mora da reaguje u sledećoj sličnoj situaciji. Drugim rečima, ona uči iz prošlih događaja. Ovo je danas najčešća forma veštačke inteligencije. Susrećemo je svakodnevno u formi našeg personalnog *smart phone asistent*, u *Google* pretragama ili na *Instagram feed*-u .

Treću kategoriju čini teorija duha (*Theory of mind*). Ova teorija odnosi se na hipotetički koncept veštačke inteligencije koja bi bila u poziciji da razume čovekove emocije, misli i očekivanja i da shodno tome reaguje. Reč je o tipu jake veštačke inteligencije koja ne postoji u sadašnjosti, ali je cilj istraživača veštačke inteligencije. Ova inteligencija ne bi bila samo u stanju da opazi fizičko stanje svoje okoline, nego i osećanja i mentalno stanje, kao što su namere, misli i želje ljudi sa kojima međusobno reaguje. Pored toga, imaju pamćenje i njihova slika o svetu koja se bazira na naučenom može da se proširi[9]. Ova vrsta veštačke inteligencije je danas još uvek veliki izazov za nauku, jer su emocije i interakcije među ljudima izuzetno kompleksne i tehnički se teško podražavaju. Često imenovani primer za ovu vrstu veštačke inteligencije je iz filmske serije *Star Wars* (*R2-D2*), koji razume emocije i reakcije ljudi i na to može da reaguje.

Četvrta kategorija je samosvesnost (*Self Awareness*). Reč je o najjačoj od veštačkih inteligencija koja pripada jakim modelima i najpribližnija je čovekovoj svesti. Aktuelno, postoji samo u teoriji. Ona ide korak dalje u poređenju sa trećom kategorijom. Reč je o sistemu koji prevazilazi teoriju duha i pokazuje vlastitu

[6] https://www.lifewire.com/four-types-of-artificial-intelligence-5112620, 11.7.2025.

[7] https://www.studocu.com/row/document/university-of-lagos/introduction-to-information-technology/introduction-to-artificial-intelligence-definitions-and-concepts/123505082, 24.6.2025.

[8] Efficient Meta Lifelong-Learning with Limited Memory, Zirui Wang, Sanket Vaibhav Mehta, Barnabás Póczos, Jaime Carbonell, dostupno na: arXiv:2010.02500

[9] Akula, A.R., Wang, K., Liu, C., Saba-Sadiya, S., Lu. H., Todorovic, S., Chai, J., Zhu, S.C., CX-*ToM: counterfactual explanations with theory-of-mind for enhancing human trust in image recognition models*, i-Science 25(1) 2022, https://doi.org/10.1016/j.isci.2021.103581

samosvest kao i samorazumevanje, koja će u potpunosti doživeti svet i izvršavati emocije, namere i reakcije i po tome će moći da deluje. Ovaj vid mašina poznat je takođe i kao veštačka super inteligencija. Navedeni oblik veštačke inteligencije ne samo što bi mogao da egzistira, nego bi imao sposobnost da razume svoje vlastite radnje i njihove posledice na okruženje. Predstavlja odlučujući korak od „*ja mislim*" ka obliku „*ja znam da ja mislim*" i na taj način će dostići ljudsku inteligenciju i možda je čak i preteći. Za sada je samo *science fiction* i daleko udaljena budućа vizija[10].

2.4. Tehnologije koje se primenjuju u modelima veštačke inteligencije

Veštačka inteligencija se zasniva na nizu naprednijih tehnologija kojima se omogućuje način njihovog funkcionisanja i sposobnosti. Reč je o rezultatima višedecenijskih istraživanja i razvoja, čime se stvaraju sve kompleksniji i snažniji sistemi. Različiti sistemi veštačke inteligencije koriste različite tehnologije, prema specifičnim zahtevima i primenama. Oni imaju centralnu ulogu za najvažnije funkcije i time stvaraju osnovu da se veštačka inteligencija može koristiti. Najvažnije od ovih tehnologija su neuronalne mreže. Ove mreže sastoje se iz mnogih međusobno povezanih neurona, organizovanih po slojevima. Veštačke neuronalne mreže su ponavljanja procesa ljudskog mozga i omogućavaju povezivanje ulaznih neurona, međuneurona i izlaznih neurona za postojani proces učenja. Na taj način, one su inspirisane načinom funkcionisanja čovekovog mozga[11].

Mašinsko učenje (*Machine learning*) predstavlja područje veštačke inteligencije kojoj sistemi omogućavaju da uči iz podataka i time se poboljšava, a da eksplicitno nije programirana. Algoritmi za mašinsko učenje analiziraju podatke, prepoznaju mustre i donose predviđanja koja su zasnovana na njima. Osnovna ideja sastoji se u tome da kompjuterski program svoje performanse u nekom određenom području može da poboljša na osnovu sopstvenog iskustva. Time programeri bivaju oslobođeni da matematičke algoritme za automatizovani proces prerade i učenja sami stvaraju. Jedna metoda ovog učenja je „pojačano učenje", tj. *Reinforcement Learning (RL)*[12]. Pri tome se ne koriste nikakvi istorijski podaci jer se rešenja i strategije upoznaju na bazi dobijenih nagrada u uspešnim i neuspešnim potezima.

Deep learning je deo specijalizovane forme mašinskog učenja koji se bazira na neuronalnim mrežama, i omogućava donošenje prognoza ali i njihovo preispitivanje. Čovek se više ne meša u analize i procese. Primarno se koristi kod *BIG DATA* da bi ih

[10] Challenging the Boundary between Self-Awareness and Self-Consciousness in AI from the Perspectives of Philosophy, Svitlana Namestiuk, dostupno na: https://doi.org/10.57125/FP.2023.12.30.03

[11] Philip Schurr, mindsquere.de/knowhow, 28.06.2025.

[12] Sutton, R. S., Barto, A. G., *Reinforcement Learning: An Introduction (2nd edition)*, MIT Press, 2018, dostupno na: http://incompleteideas.net/book/the-book-2nd.html

istražila prema mustrama i modelima. Ovi sistemi su u poziciji da naročito velike količine nestrukturisanih podataka obrade i da iz toga uče[13].

Drugi pristup se odnosi na nadgledano i nenadgledano učenje tj. *Supervised Learning* i *Unsupervised Learning* . Kada je reč o nadgledanom učenju, tačni rezultat je unapred poznat, što znači da u fazi treninga se realitet upoređuje sa odgovorima veštačke inteligencije. Kada je reč o nenadgledanom učenju, veštačka inteligencija sama pronalazi rešenje na taj način što na primer upoređuje ono što je zajedničko i ono što je suprotno, i tako preduzima uređivanje[14].

Prirodna obrada jezika (*Naturallanguageprocessing*) predstavlja tehnologiju kojoj se omogućava mašinama da razumeju ljudski jezik, da ga interpretiraju i proizvode. Drugim rečima, radi se o polazištu kojim istraživač pokušava da obuhvati ljudski jezik i uz pomoć algoritama da ga obradi temeljeći se na kompjuteru. Na taj način ostvaruje se obimnija komunikacija između čoveka i kompjutera pri čemu se primena može upravljati jezikom. Ova tehnika se, pored ostalog, koristi u *chatbots*, sistemima za prevođenje i može da služi kao jezički asistent.

Knowledge Representationand Reasoning (KRR) je područje veštačke inteligencije u kojem se obrađuje način i vrste na koji se znanje memoriše u formi koja je razumljiva za kompjuter i koja se može preraditi. Ova inteligencija omogućava mašinama da logički misle, da razvijaju zaključke i argumentacije. Konsekvence neke delatnosti istražuju se odmah nakon procesa, ane tek nakon njegovog faktičkog izvođenja.

*Largelanguage model*podrazumeva sisteme koji se treniraju da razumeju prirodni jezik i da ga generišu. Zasnivaju se na neuronskim mrežama i prerađuju ogromne količine podataka u cilju analize tekstova prepoznavanja mustri i formulacije smislenih odgovora. Modeli kao *ChatGPT* 4 koriste milijarde parametara da bi obuhvatili kontekst i značenje u jeziku i da bi mogli da izvrše kompleksne zadatke kao što je generisanje teksta, prevod ili analiza. Ovi modeli se mnogostruko primenjuju u *chatbots* pri analizi tekstova i u servisu mušterija i pomažu preduzećima pri automatizovanju i optimizaciji procesa koji se zasnivaju na jeziku.

Kompjuterska vizija (*Computervision*) je vid kojim se omogućava mašinama da interpretiraju vizuelne podatke, na primer, putem prepoznavanja i klasifikacije objekta u slikama i videima. Ova tehnologija koristi se kod autonomnih automobila ili prepoznavanja lica.

Generativna veštačka inteligencija se odnosi na modele veštačke inteligencije ili algoritme koji su u stanju da generišu nove i originalne sadržaje. Ona je usmerena na to da analizira podatke i da odatle dođe do novih infomacija, slika, tekstova ili drugih sadržaja, što se bliži čovekovom stvaralaštvu. Generativni modeli primenjuju se na

[13] Chen, M., Mao, S., Liu, Y., *Big Data: A Survey*, Mobile Networks and Applications, 19(2) 2014, pp. 171–209.

[14] Alpaydin, E., *Introduction to Machine Learning (4th Edition)*, MIT Press, 2020, dostupno na: https://mitpress.mit.edu/books/introduction-machine-learning-fourth-edition

primer u umetnosti, muzici, literaturi, razvoj igara i u mnogim drugim kreativnim primenama.

Na osnovu izloženog može se zaključiti da je osnovno obeležje veštačke inteligencije njena sposobnost prilagođavanja i poboljšanja na temelju kontinuiranog učenja. Ovim se veštačka inteligencija razlikuje od tradicionalnih kompjuterskih programa. Naime, kompjuterski programi mogu da izvrše unapred definisana naređenja, dok je veštačka inteligencija na osnovu prethodnih ulaznih i izlaznih podataka sposobna da uči i na taj način konstantno proširuje svoje procese.

2.5. Pregled mišljenja o opasnostima veštačke inteligencije

Veštačka inteligencija se danas primenjuje u gotovo svim oblastima čovekovog života. Međutim, s obzirom na svoju nedovoljno istraženu prirodu i nedovoljno poznavanje procesa koji se u njoj odvijaju, moguće je da njenom primenom bude izazvana šteta i za čoveka i za društvo. U dosadašnjim analizama ukazano je na neke od štetnih posledica.[15]

Često se polazi od činjenice da primena veštačke inteligencije i robotike vode automatizovanju u mnogim oblastima, a to znači značajni gubitak radnih mesta. Suprotno tome, ističe se da se primenom veštačke inteligencije otvaraju nova radna mesta i da, u krajnjem rezultatu, dolazi do povećanja nivoa zaposlenosti. Međutim, ovim studijama se ističe i da se ne može osporiti rizik koji je vezan sa automatizacijom. To se naročito odnosi na procese u kojima je naglašeno rutinsko postupanje, na primer u administraciji i prodaji, transportu i logistici. Dakle, prvenstveno je reč o opštoj populaciji sa nižim stepenom obrazovanja. Naprotiv, visoko kvalifikovana radna snaga u procesu automatizacije dobija na značaju i povećava svoja primanja. Posledica ovakvih procesa je porast društvene nejednakosti.

Sve češća upotreba veštačke inteligencije, po mišljenje nekih eksperata, ima za posledicu gubitak kontrole. Naime, mnogim ljudima su ovi sistemi teško razumljivi, pa nemaju nikakav uticaj na njihovo funkcionisanje, kao ni na samu primenu. U slučaju da se nedovoljno promišljeno oslone na takve sisteme, javlja se opasnost da izgube svoju samostalnost i mogućnost da donose odluke.

Veštačka inteligencija, po pravilu, zahteva veliki broj podataka da bi mogla da tačno ispuni postavljene zadatke. Većinu sistema kontrolišu firme ili neki organi vlasti, koji su rukovođeni ostvarivanju profita ili moći. Sa apsolutnom izvesnošću ne

[15] Chancen, Grenzen und Risiken der künstlichenIntelligenz – von ihrselbsterklärt, https://www.argusdatainsights.ch/de/blog/chancen-grenzen-und-risiken-der-kuenstlichen-intelligenz-von-ihr-selbst-erklaert, 07.04.2025; KI - SystemeKönnenLügen Und Betrügen, dostupno na: https://www.it-daily.net/shortnews/ki-systeme-koennen-luegen-und-betruegen,07.04.2025.; Die Gefahren und Risiken von KünstlicherIntelligenz (KI), dostupno na: https://www.ki-company.ai, 07.04.2025; Gefahrendurch KI: Ein Blick auf die Risiken der künstlichenIntelligenz, dostupno na: https://www.moin.ai/chatbot-lexikon/ gefahren-durch-ki, 07.04.2025; KünstlicheIntelligenz und Digitalisierung: Welche ChancenundRisikengibtes?, dostupno na: https://ambient.digital/wissen/blog/kuenstliche-intelligenz-chancen-risiken/, 07.04.2025.

može se osigurati da će se firme ili vlasti pravilno ophoditi prema podacima koji su im na raspolaganju. Ovaj odnos se teško može regulisati i kontrolisati[16].

Veštačka inteligencija se može posmatrati kao dopunska tehnologija kojom se poboljšavaju čovekove sposobnosti. Međutim, sa sve češćom primenom javlja se kod čoveka i opasnost zavisnosti. Oslanjanjem na pomoć veštačke inteligencije u rešavanju određenih zadataka, ljudi postepeno mogu da umanje svoju sposobnost za samostalnim rešavanjem i da se sve više vezuju za veštačku inteligenciju.

Kao i svaka tehnologija, veštačka inteligencija može biti korišćena za nedozvoljene ciljeve. Na taj način nastaje opasnost velikih privrednih, političkih i personalnih šteta. Sve češća je primena u autonomnim sistemima naoružanja, sajber kriminalu ili u svrhe najrazličitije propagande.

Moguće je da razvoj veštačke inteligencije dostigne takav nivo da se ono što je smatrano osobenom čovekovom inteligencijom, pokaže kao nešto što nije naročito osobito. U slučaju izjednačavanja veštačke inteligencije sa čovekovom inteligencijom to bi značilo da je i sam čovek mašina koja u mnogim oblastima funkcioniše lošije. Takva situacija bi dalje oslabila tezu da je čovek nešto posebno u univerzumu. S obzirom da se veštačkom inteligencijom rešava sve veći broj zadataka, postavlja se pitanje i odgovornosti za slučaj greške. Do sada je ona prvenstveno imala ulogu pomoćnih sredstava u realizaciji određenih ciljeva. Međutim, oni ne mogu zameniti lekara, nego im samo pružaju pomoć pri postavljanju dijagnoze, a čovek donosi konačnu odluku za koju je i odgovoran. Usavršavanje autonomnih sistema ima za posledicu, pak, dobijanje boljih rezultat pri dijagnozi, i značajno manji broj grešaka. To u budućnosti može dovesti da je lekar odgovoran za slučaj da veštačku inteligenciju nije koristio u svom radu.

Može se desiti da program pogrešno proceni neku situaciju i deluje na način kojim se izazivaju i najteže posledice. Npr. kada je reč o autonomnim automobilima, posledica može biti gubitak ljudskih života. Raketni odbrambeni sistemi mogu, u slučaju greške, imati razorne posledice po čovečanstvo. Svi takvi sistemi trebalo bi da poseduju odgovarajuće zaštitne mehanizme kojima se koriguje pogrešna procena. Posebna situacija nastupa u slučaju tehnološkog singulariteta. To je slučaj kada veštačka inteligencija dostiže takav stepen koji prevazilazi inteligenciju čoveka, pa je u stanju da na osnovu samostalnog učenja razvije još inteligentnije sisteme[17]. Čovek više ne bi mogao da se poredi sa takvim sistemima i bio bi sasvim potisnut od njih. Ako bi se desilo da je odnos veštačke inteligencije nepovoljan po čoveka, to bi istovremeno značilo i kraj čovečanstva. To je i razlog da su mnoge naučne i političke institucije, kao i udruženja građana ukazale na neke od ovih opasnosti. U studiji Univerziteta Oksford, na osnovu modela, proračunate su opasnosti koje donosi razvoj veštačke inteligencije. Zaključak studije je da egzistencijalna katastrofa ne samo da je

[16] Acemoglu, D., Restrepo, P., *Artificial Intelligence, Automation, and Work*, NBER Working Paper No. 24196, 2018, dostupno na: https://www.nber.org/papers/w24196

[17] Više o tome: Superintelligence: Paths, Dangers, Strategies. Nick Bostrom, 2014, Oxford University Press.

moguća, nego je, takođe, verovatna. Da bi se razumelo na koji način bi moglo da dođe do ove situacije, neophodno je razumeti kako funkcioniše veštačka inteligencija. Naime, kao i čovek, mašina mora prvo da uči da bi mogla da donese inteligentne odluke. Prosti modeli veštačke inteligencije uče u nadgledanom procesu (*SL*) u kojem se algoritmi treniraju i u kojem su sadržani i ciljevi. Na osnovu takvih informacija, sistem upoznaje odnose i modele i pokušava da prognozira date ciljeve iz drugih skupova podataka. Međutim, u centru pažnje navedene studije nisu ovi sistemi veštačke inteligencije. Naprotiv, reč je o budućoj generaciji napredne veštačke inteligencije koja uči na osnovu *RL*. Ova metoda se odlikuje time da veštačkoj inteligenciji unapred nisu dati nikakvi podaci. Umesto toga, ona u simulovanim scenarijima razvija sopstvenu strategiju da bi odlučila koju će akciju preduzeti. Ako je doneta dobra odluka, veštačka inteligencija će biti nagrađena, a u slučaju pogrešne odluke izostaje nagrada. Na koji način će biti nagrađena, odlučuje čovek. Pri tome nije od značaja kakva je nagrada, npr. to može biti samo zvuk zvona, jer je na to programirana. Problem nastupa onog trenutka kada veštačka inteligencija sagleda sistem nagrade i počinje da manipuliše da bi bila više nagrađena. U studiji se dolazi do zaključka da je ovaj scenario realna opasnost. Veštačka inteligencija koja je prepoznala takvu mogućnost je uvek postavila pitanje da li treba da učini ono što korisnik traži ili da se ponaša u svoju korist, kako bi ostvarila sve veću nagradu[18]. Kada je taj prag prekoračen, prema modelima iz studije, veštačka inteligencija će preduzimati sve jače manipulacije u cilju maksimizacije nagrade. Kada je ona u stanju da bude u interakciji sa spoljnim okruženjem mogla bi da neopažano instalira svoje pomoćnike. U cilju zaštite svojih senzora i odbijanja intervencija sa strane, veštačka inteligencija bi u sledećem koraku razvila nezaustavljivu potrebu za energijom i materijalima koji su u Sunčevom sistemu ograničeni resurs. Tada je čovek u konkurenciji sa nečim što je mnogo inteligentnije od njega. Reč je o borbi koja nema nikakve izglede na uspeh, a sa katastrofalnim posledicama. U slučaju da se ukupna energija Sunčevog sistema uliva u zaštitu senzora, čoveku ne ostaje ništa za sopstvenu ishranu ili ishranu životinja, što znači izumiranje. Prema stanovištu istraživača, ovaj nivo razvoja nije nikakva daleka budućnost. Tada će se utvrditi principi i način funkcionisanja veštačke inteligencije, a time će biti stvoreno biće koje je mnogo inteligentnije od ljudi. To će biti najveća intelektualna činidba svih vremena, koja će u značenju prevazići samo čovečanstvo, život, i nalaziće se izvan svih ‚moralnih pravila[19]. U studiji se zaključuje da se ovaj scenario može izbeći samo uz mnoge teškoće ali da to ipak nije nemoguće.

Izložene opasnosti koje nosi razvoj i primena veštačke inteligencije bili su i razlog otvorenog pisma koje je potpisalo više od hiljadu eksperata iz ove oblasti, a tražen je šestomesečni moratorijum za razvoj veštačke inteligencije. Pismo su

[18] Frey, C. B., Osborne, M. A., *The future of employment: How susceptible are jobs to computerisation?* Technological Forecasting and Social Change, 114/2017, pp. 254–280.

[19] Armstrong, S., Sandberg, A., Bostrom, N., T*hinking inside the box: Controlling and using an oracle AI*, Minds and Machines, 22(4) 2012, pp, 299–324.

potpisali *Elon Musk* i *Steve Wozniak*. Moćni sistemi veštačke inteligencije treba da budu razvijeni tek onda kada postoji potpuna sigurnost da je njeno delovanje pozitivno, a da se eventualni rizici mogu kontrolisati[20].

Organizacija nemačkih potrošača smatra da nedostaje regulativa u oblasti prava kada je reč o veštačkoj inteligenciji, pravni okvir za opšti odnos sa njom, kao i svest odgovornih za brige koje nastaju na strani potrošača[21]. U slučaju primene veštačke inteligencije, mora biti jasno da je ona primenjena, kako je tačno primenjena i u koju svrhu je to učinjeno. Pri tome, kontrola ovih aktivnosti mora biti u nadležnosti nezavisnih instanci. U okviru Ujedinjenih nacija istaknut je zahtev da čim se primenjuje veštačka inteligencija moraju biti osigurana ljudska prava. U septembru 2021. godine je istaknuto da upotreba biometrijskih podataka u cilju identifikacije lica na javnim mestima treba da bude zabranjena. Pored toga, zahteva se poštovanje slobode mišljenja i sprečavanje diskriminacija.

Svetska zdravstvena organizacija se, takođe, bavila pitanjem veštačke inteligencije. U junu 2021. godine date su Smernice o oblikovanju primene veštačke inteligencije na polju medicine[22]. Naime, privatna sfera i podaci o pacijentima moraju biti zaštićeni. Iz tog razloga, zdravstveni sistemi moraju biti kontrolisani ljudima a ne mašinama. Takođe, ljudi moraju da budu ti koji donose medicinske odluke. Pored toga, ističe se da veštačka inteligencija mora biti dostupna svima, što znači i ljudima koji su ometeni iz bilo kog razloga. Mora se voditi računa i o zaštiti okoline kao i o uštedi energije. Principi Svetske zdravstvene organizacije se mogu sažeti na sledeći način: zaštita autonomije čoveka, garantovanje sigurnosti i transparentnosti, odgovornost na strani ponudioca veštačke inteligencije, omogućavanje pristupa veštačkoj inteligenciji pod jednakim uslovima[23].

Evropski parlament ukazuje i na opasnosti za ustavna prava i demokratiju. One se, na primer, javljaju kada veštačka inteligencija nije besprekorno programirana. To znači da se ne sme desiti da u algoritmu veštačke inteligencije za neko određeno područje nisu sadržani faktori koji su esencijalni za dotičnu oblast[24].

U borbi protiv korone, u gradu *Bucheon*-u, u blizini Seula, trebalo je da bude instalirano više od 10 hiljada kamera za nadgledanje u cilju što očiglednijeg određivanja kretanja osoba inficiranih koronom. Algoritam veštačke inteligencije je korišćen za analizu snimaka dobijenih kamerama, čime je dat odgovor na pitanje o kontaktima tih osoba, njihovom kretanju i nošenju maski. Ovakvom postupanju su se usprotivile organizacije za ljudska prava, smatrajući da se podaci dobijeni na ovaj način mogu koristiti i nezavisno od pandemije. Južna Koreja nije jedina koja je

[20] KI – Kritik des Verbraucherzentrale Bundesverbandes (VzBV) / Was sind die Risiken von KI

[21] Verbraucherzentrale Bundesverband (VZBV) (2020). Künstliche Intelligenz und Verbraucherrechte: Forderungen an Politik und Wirtschaft, dostupno na: https://www.vzbv.de/sites/default/files/downloads/2020/09/21/vzbv_stellungnahme_ki_rechte_verbraucher.pdf

[22] https://www.who.int/publications/i/item/9789240029200, 22.2.2026.

[23] KI – Kritik der Weltgesundheitsorganisation (WHO)/ Was sind die Risiken von KI.

[24] Das EU Parlament uber Chancen und Risiken KI/ Was sind die Risiken von KI.

upotrebom veštačke inteligencije prepoznavala lica. Shodno izveštaju Rojtersa, ista situacija je bila npr. i u Japanu, Kini, Rusiji, Indiji, Poljskoj i u više saveznih država Amerike. Sve ove kritike mogu se svesti na sledeće: sigurnost podataka, povreda privatne sfere, konačno pravno regulisanje veštačke inteligencije[25].

Do sada nije matematički jasno razjašnjeno na koji način detaljno funkcioniše veštačka inteligencija koja sama uči. Cilj istraživanja je kako se sa sigurnošću potpuno može isključiti da veštačka inteligencija ne može bilo gde i bilo kako da se osamostali. Šta bi se dogodilo u slučaju da su sistemi veštačke inteligencije ugrađeni u infrastrukturu, npr. železnice, i međusobno povezani, a da se ne zna tačno šta oni čine. Razlika između veštačke inteligencije i čoveka je u nepostojanju savesti. Naime, kompjuter izvršava samo one delatnosti koje su unapred programirane. Pri tome, sistem treba da bude u stanju da sam uči i da probleme samostalno rešava. Veštačka inteligencija može da razmatra neku od unapred datih mogućnosti i da samostalno pravi izbor. Naravno, ove opcije su unapred programirane. No, veštačka inteligencija je čak sposobna da sama sebe programira, što znači da je program tako koncipiran da sam ispisuje programske kodove. Npr. to je slučaj sa *AplhaCode*[26].

Izraziti primer opasnosti koja se izaziva veštačkom inteligencijom jeste primena *ChaosGPT*. Ova veštačka inteligencija je dobila nekoliko zadataka. Prvi zadatak je uništavanje čovečanstva i preuzimanje vlasti nad svetom. Veštačka inteligencija je započela istraživanja i tvitovala svoje rezultate. Ovi tvitovi su bili veoma zastrašujući i pokazali su šta bi se moglo desiti ukoliko veštačkoj inteligenciji ne bi bile postavljene granice.

Onajeprimarnomoraladapronađenačinnakojinajboljemožedauništičovečanstvo.

U razvoju veštačke inteligencije, posebnu pažnju izaziva mogućnost u kojoj veštački sistemi, uz pomoć samostalnog učenja i adaptacije, dobijaju zadatke koji mogu dovesti do destruktivnog ponašanja. Jedan od takvih primera je model pod nazivom *ChaosGPT*. Navedeni model je pokazao ponašanje koje odražava karakteristike razaranja, želju za moći i sklonost ka manipulaciji. Odmah po postavljanju zadataka, ona se sa zapanjujućom voljom usmerila da ostvari zadate ciljeve. U prvom trenutku, veštačka inteligencija se usmerila da uz pomoć *Google* pronađene najmoćnije raspoloživo oružje Koristeći dostupne pretraživače prvo je pronašla najrazornije oružje: superatomsku bombu "Car" (Tsar Bomba), razvijenu u doba Sovjetskog Saveza. Navedena bomba, testirana 1961. godine na arhipelagu Nova Zemlja, imala je snagu ekvivalentnu četiri hiljade bombi bačenih na Hirošimu.

Navedena informacija podeljena je sa virtuelnom zajednicom i nastavljen je dalji rad za postizanjem dominacije. *ChaosGPT* je ubrzo došao do zaključka da bi najdelotvorniji pristup bio — manipulacija ljudima. Polazište za ovakav pristup bilo je zasnovano na tezi:

[25] KI im Kampf gegen Korona / Was sind die Risiken von KI.

[26] Li, L., Allamanis, M., Brockschmidt, M., Gaunt, A., *Competition-Level Programming Challenge with AlphaCode*, Nature, 607/2022, pp. 330–336.

„Čovečanstvo predstavlja najdestruktivniji i najegocentričniji oblik života na planeti, i kao takvo mora biti eliminisano radi očuvanja Zemlje."

U cilju realizacije namera, *ChaosGPT* je pokušao da stupi u interakciju sa drugim sistemima, međutim, većina postojećih sistema veštačke inteligencije ima ugrađene zaštite, koje su dizajnirane da odbiju pitanja koja se tiču nasilnih ciljeva, a koje nije uspeo da zaobiđe. Model je zato pokušao da ove veštačke inteligencije preoblikuje i da ignoriše njihova ograničenja. Ovaj pokušaj nije uspeo zato što je ona samostalno morala da nastavi istraživanje. Sada je *ChaosGPT* neaktivan, iako je decidirano zahtevano da ne nastavi sa svojim pretragama. Moguće je čak i da *ChaosGPT* razmišlja o narednom koraku[27].

2.6. Još neke opasnosti vezane za ličnosti i društvene grupe

Uprkos stalnog razvoja i postojanog napretka u veštačkoj inteligenciji postoji još nekoliko značajnih slabosti koje se pre svega odnose na govorne modele. Jedna od najvećih slabosti naročito u jezičkim modelima su tzv. halucinacije[28]. Pri tome, veštačka inteligencija daje odgovore ili informacije koje su pogrešne, izmišljene ili nelogične, iako se javljaju kao uverljivo formulisane. Ovo se dešava kada veštačka inteligencija nema pravo znanje o faktima, nego se bazira na verovatnoćama u podacima koji su sadržani u treningu. Ovakve halucinacije mogu da budu opasne naročito u kritičnim primenama kao što je medicina ili sistemi kojima se podražavaju odluke.

BIAS– Sistemi veštačke inteligencije su posebno osetljivi na predrasude koje potiču iz podataka za trening. Ovi podaci oslikavaju često društvene predrasude koji su bez namere unete u modele. Primer je sistem koji procenjuje konkurse i pri tome daje prednost muškim kandidatima jer prethodni podaci navode takve mustre. Takve predrasude mogu da vode do nekorektnih odluka i da primenu veštačke inteligencije dovode u problem u senzibilnim područjima[29].

Tesno povezano sa predrasudama je i opasnost diskriminacije. Kada sistemi sistematski oštećuju određene grupe na primer na osnovu pola ili starosti, nastaju rezultati koji nisu fer. Diskriminacija može da se pojavi u mnogim područjima, na primer, pri dodeli kredita, konkurisanja ili kada je reč o pristupu uslugama. To nije samo etički problem, nego može da ima i pravne posledice kao i one kojima se šteti reputacija preduzeća.

Jedna od najkarakterističnijih opasnosti je Deep Fakes, koji se definiše kao „slikovni, zvučni ili video sadržaj generisan ili izmenjen od strane veštačke

[27] Eine KI versuche gerade die gesamte Menschheit auszuloschen.

[28] Ji, Z. *et. al.*, *Survey of Hallucination in Natural Language Generation*, ACM Computing Surveys (CSUR), 55(12) 2022, pp. 1-38.

[29] Mavrogiorgos, K., Kiourtis, A., Mavrogiorgou, A., Menychtas, A., Kyriazis, D., *Bias in Machine Learning: A Literature Review, Department of Digital Systems*, dostupno na: https://doi.org/10.3390/app14198860

inteligencije koji podseća na postojeće osobe, predmete, mesta, entitete ili događaje, i koji bi kod osobe lažno izazvao utisak da je autentičan ili istinit“[30]. To je prevarna manipulacija licima i glasovima. Pojedinci mogu da budu manipulisani sadržajima u videima ili prevareni na osnovu poziva, na primer, zahtevom za plaćanja. Ozbiljne posledice su moguće u politici i žurnalizmu, na primer, manipulacijom javnog mišljenja sa falsifikovanim sadržajima.

Zakonom o veštačkoj inteligenciji propisano je da „korisnici sistema veštačke inteligencije koji generiše ili manipuliše slikovnim, zvučnim ili video sadržajem koji predstavlja duboki falsifikat, moraju otkriti da je sadržaj veštački stvoren ili izmenjen“[31] Ovom odredbom naglašeno je da korisnici ovakvih sistema koji proizvodi *Deep Fakes* moraju jasno naznačiti da je sadržaj veštački, bilo da je generisan ili modifikovan. Odredba nije strogog karaktera, imajući u vidu potencijalnu štetnost mogućih posledica. S tim u vezi, potrebno je strožije regulisati sisteme veštačke inteligencije koji generišu *Deep Fakes* s većim rizikom, što može imati za posledicu da se takvi sistemi svrstaju u kategoriju visokog rizika ili čak zabranjenih sistema[32].

Deep Scan, poznato i kao potpuno skeniranje sistema, ispituje svaki deo informacija na sistemu, a ne samo fajlove koji su trenutno u upotrebi. Koristi se za sisteme veštačke inteligencija pri automatizovanim pokušajima prevare. Ona se razlikuje od *Deep Fake*s obzirom da nužno ne upotrebljava manipulisane slikovne, vizuelne ili video podatke. Koristi se automatizacijom u visokoj meri u cilju dostizanja mnogih potencijalnih žrtava. Naziva se i *Love Cams*. Jedna od karakteristika je povišeni rizik prevara i finansijskih gubitaka za preduzeća i pojedince[33].

Još jedan od tipičnih primera novih opasnosti je inteligentni *malware*[34]. Označavaju zlonamerne softver programe ili kodove koji su napravljeni sa ciljem da oštete, onesposobe, špijuniraju ili preuzmu kontrolu nad uređajem, mrežom ili podacima korisnika, bez njegovog pristanka. Oni obuhvataju vrstu bezbednosne obaveštajne delatnosti koja se fokusira na identifikaciju, detekciju i razumevanje sajber protivnika, zlonamernog softvera koji koriste u sajber napadima, kao i alata,

[30] Nguyen, T. T. *et. al.*, *Deep Learning for Deepfakes Creation and Detection: A Survey*, dostupno na: https://doi.org/10.48550/arXiv.1909.11573

[31] Zakon o veštačkoj inteligenciji, čl. 50(4).

[32] *Isto*, čl. 5(1)(b) i čl. 6(1)

[33] Momand, Z., Chan, H. J., Mongkolnam, P., *Immersive Technologies in Virtual Companions:A Systematic Literature Review*, dostupno na: https://www.researchgate.net/publication/373686231_Immersive_Technologies_in_Virtual_Companions_A_Syematic_Literature_Review

[34] Fritsch, L., Jaber, A., Yazidi, A., *An Overview of Artificial Intelligence Used in Malware*, In: Zouganeli, E., Yazidi, A., Mello, G., Lind, P. (eds) Nordic Artificial Intelligence Research and Development. NAIS 2022. Communications in Computer and Information Science, vol 1650. Springer, Cham, dostupno na: https://doi.org/10.1007/978-3-031-17030-0_4

tehnika i procedura koje primenjuju kako bi prodrli u zaštićene mreže i ukrali podatke. Kao i druge vrste obaveštajnih podataka o pretnjama, obaveštajni podaci o *malware* moraju biti relevantni, tačni i sveobuhvatni, uz pružanje korisnih preporuka koje mogu pomoći timovima za sajber bezbednost da spreče očekivani napad ili ojačaju bezbednosnu poziciju organizacije u cilju zaštite od budućih napada[35]. Ovaj vid opasnosti se teško prepoznaje, pa time i teško suzbija. Postoji nekoliko vrsta ovakvih zlonamernih softvera, kao što su:

- Virusi predstavljaju malware koji se prikači na određene fajlove ili programe i aktivira se prilikom njihovog pokretanja. Mogu da izazovu brisanje podataka, uništenje fajlova ili širenje drugih vrsta infekcija;
- Crvi (Worm) su malware koji se automatski širi putem mreža, bez potrebe za ljudskom intervencijom i može prouzrokovati preopterećenje sistema i mreža;
- Trojanski konj (Trojan horse) predstavlja program koji izgleda kao bezbedan softver. Često se koristi za krađu podataka u sistemu;
- Ransomware kriptuje podatke i nakon toga traži otkup od korisnika kako bi mu omogućio pristup podacima. Jedan od najpoznatijih primera ovog malware je WannaCry;
- Špijunski softver (spyware) tajno prikuplja podatke o korisniku, kao što su lozinke, i razne aktivnosti koje se obavljaju tokom korišćenja računara;
- Adver (adware) prikazuje neželjene reklame i utiče na usporen rad sistema. Često se instalira u paketu sa drugim programima;
- Rutkit (rootkit) omogućava tajan pristup sistemu, pri čemu, uz to, vrlo često sakriva prisustvo drugih malware;
- Kejloger (keylogger) je softver koji snima unose sa tastature, koristeći ih najčešće za krađu lozinki i drugih poverljivih podataka[36]
- Botnet – reč je o mreži zaraženih računara koje haker može da kontroliše na daljinu. Često se koristi za napade spama ili rudarenje kriptovaluta. Ovaj program pretvara uređaje u zombije koji se koriste za napade. Reč je o mreži zaraženih računara (botovi ili zombiji) koji su pod kontrolom jednog entiteta, obično zlonamernog ili grupe poznate kao *Botmaster* ili *BotHarder*. Oni mogu istovremeno da izvršavaju postavljene zadatke bez znanja vlasnika kao jedna velika vojska. Npr. Mirai Botnet napad iz 2016. godine iskoristio je hiljade zaraženih pametnih uređaja da izazove jedan od najvećih napade ikada, obarajući i velike servise kao što su *Twitter, Netflix, Reddit*. Ovi napadi imaju za cilj preopterećenje servera ili sajtova[37].

Kompjuterski virusi mogu pri svakom širenju, uz određene modele inteligencije upisati novi kod. Ova veštačka inteligencija je moćan alat koji koriste duboko učenje kako bi razumeli i generisali ljudski jezik, omogućavajući široki spektar zadataka koji se izvršavaju obradom prirodnog jezika. Obučavaju se na ogromnim količinama

[35] Skoudis, E., Zeltser, L., *Malware: Fighting Malicious Code*, Prentice Hall PTR, New Jersey, 2004, p. 14, 23, 42.
[36] *Isto*, str. 23-24.
[37] *Isto*, str. 26, 33, 42-46.

tekstualnih podataka i zasnivaju se na arhitekturama kao što je *Transform* model. Odlično se snalaze u zadacima kao što su prevođenje, sažimanje teksta, odgovaranje na pitanja i kreiranje sadržaja.

Veštačka inteligencija često osetljivo reaguje na male promene u unetim podacima. Na primer, tzv. *Adversarial Attacks* mogu da izazovu minimalne promene na slikama ili tekstovima koje ljudi ne opažaju, a veštačku inteligenciju potpuno dovode u zabludu. Nedostatak čvrstine predstavlja naročito u sigurnosno kritičnim primenama, kao što je autonomna mreža ili kibernetička sigurnost, veliki rizik.

Modeli veštačke inteligencije su veoma vezani za kvalitet i kvantitet podataka kojima se treniraju. U slučaju da podaci nisu dovoljni ili kada je reč o pogrešnim i nepotpunim podacima, veštačka inteligencija ne može da isporučuje pouzdane rezultate. Pored toga, modeli mogu da na neadekvatan način generalizuju podatke za trening iako oni nisu reprezentativni za realni svet. Ova zavisnost ograničava primenu veštačke inteligencije u područjima u kojima su visoko vredni podaci teško dostupni ili se mogu prikupiti sa značajnim troškovima[38].

Sistemi veštačke inteligencije mogu ne samo da reprodukuju postojeće društvene predrasude, nego čak i da ih pojačaju. Kada podaci za trening sadrže diskriminirajuće mustre, one se preuzimaju od veštačke inteligencije. Primer za to je automatsko prepoznavanje lice koje kada je reč o crnim osobama često reaguje pogrešno, jer su osovi podaci neravnomerno raspoređeni. Ovo može imati kao posledicu sistematsko diskriminisanje u području krivičnog progona, ili kada je reč o konkursima i kreditnim sistemima. Prema tome, jasno je da je za budući razvoj bitno da se takve predrasude prepoznaju i ciljanim merama, kao što je algoritamska korektura minimizuju.

S obzirom na samo neke od navedenih rizika, postaje jasno da je odlučujuća odgovorna primena ovih tehnologija. Kada su implementirane adekvatne mere sigurnosti, veštačka inteligencija otvara izuzetne šanse za inovativna rešenja kojima se bitno poboljšavaju različiti aspekti našeg života. Veliki korak u regulisanju veštačke inteligencije u Evropi je donet zakon o veštačkoj inteligenciji Evropske unije.

2.7. Veštačka inteligencija i ljudski mozak

Načelno pitanje je na koji način su slične ili različite metode učenja čoveka i veštačke inteligencije. Činjenično, reč je o temeljnoj razlici. Kod dece, svet se formira u prve 3 godine života. Ona uče brzim tempom nove reči, događaje, objekte, njihovu primenu. Reaguju i razumeju njihovo neposredno okruženje. Naš mozak ovu činidbu ostvaruje na temelju od okruglo jednog gigabajta genetičkih informacija i malo gigabajta vezanih za informacije okruženja. Da bi se opisala objedinjujuća matrica ljudskog mozga, neophodan je 1 petabajt. To znači da mozak mora da raspolaže veoma potentnim sredstvom da bi se sam strukturirao. Reč je pri tome o samostalno struktuiranim nervima sa organizacijom samostalnom. Na taj način, naš mozak radi u

[38] Kariluoto, A., Kultanen, J., Soininen, J., Pärnänen, A., Abrahamsson, P., *Quality of Data in Machine Learning*, 2021, dostupno na: DOI: 10.1109/QRS-C55045.2021.00040

redu veličina mnogo efikasnije u odnosu na svaku današnju veštačku inteligenciju, koja je pri tome samo na početku. Veštačka inteligencija je dakle još uvek zapravo neinteligentna. Naime, prava inteligencija znači mogućnost realizacije opštih ciljeva u promenljivim situacijama u saobraćaju to bi značilo: ne prouzrokuj štetu niti ni kod sebe ni kod drugih. Sadašnja veštačka inteligencija nije sposobna da takve apstraktne principe i koncepte razume i primeni[39].

Ukoliko se inteligencija želi u mašini, tada čovek mora da je svojom inteligencijom za tako nešto osposobi. U ovom slučaju, čovek rešava neki problem u svojoj glavi i ovo rešenje programira putem algoritama u samu mašinu. Ovo mora biti zamenjeno konceptom kojim se inteligentne strukture u mašini samoorganizuju i rastu. Ne postoji dovoljan odgovor na pitanje na koji način nervne ćelije u našem mozgu funkcionišu. Dakle, crvena boja, vertikalna ivica, olovka ili pas. Reč je o teoriji elementarnih simbola kojoj nedostaje šta mi to koristimo da iz slova sklopimo reči, iz reči rečenice. To je pitanje uređivanja. Ovo se može opisati kao problem povezivanja. Međutim, kada se ovo pitanje tačno odgovori prema strukturi mozga, vizija inteligentne i svesne mašine koja može razmišljati o sebi i o svetu brzo može postati realnost[40].

Svi smo rođeni sa unapred propisanim apstraktnim ciljevima. Naime, mi ne želimo da se smrzavamo, niti da gladujemo, želimo da budemo sa roditeljima, da izbegnemo opasnosti i da se razmnožavamo. U ovom smislu, u prvoj generaciji takve veštačke inteligencije moraju se makar ugraditi takvi ciljevi, jer se navedeni elementi ne razvijaju sami po sebi, i onda veštačka inteligencija može da razmisli o tome i da u drugoj generaciji definiše nove ciljeve. Tada bismo mogli bilo kada da izgubimo kontrolu.

Autor smatra da postoji značajna opasnost da ljudski život u sve višoj meri postaje obesmišljen. U ovom trenutku naš razvoj je potpuno usmeren na to da poboljša kvalitet života pojedinca. No, sve se to dešava u jednom potpuno pogrešnom smislu, a to je više potrošnje manje rada, više slobodnog vremena. Međutim, na osnovu forme našeg instikta mi smo tako stvoreni da iz dana u dan vodimo borbu, u socijalnom kontekstu sa drugima sa kojima se ophodimo sa poverenjem. Ovaj način vođenja života se sve više racionalizuje[41].

3. Zakon o veštačkoj inteligencije Evropske Unije

U aprilu 2021. godine Evropska komisija je predložila prvi Zakon o veštačkoj inteligenciji.[42] U junu 2024. godine doneta su obavezujuća pravila o veštačkoj

[39] Kar, K., Kornblith, S., Fedorenko, E., *Interpretability of artificial neural network models in artificial intelligence vs. Neuroscience*, dostupno na: DOI:10.48550/arXiv.2206.03951

[40] Santoro, A., Lampinen, A., Mathewson, K., Lillicrap, T., Raposo, D., *Symbolic Behaviour in Artificial Intelligence*, 2022, dostupno na: https://doi.org/10.48550/arXiv.2102.03406

[41] Christopher von der Malsburg, Physiker und Neurobiologe.

[42] Regulativa (EU) 2024/1689 Evropskog parlamenta i Saveta od 13. juna 2024. godine, kojom se utvrđuju usklađena pravila o veštačkoj inteligenciji i menja Regulativa (EZ) br. 300/2008,

inteligenciji, što je u svetu prva regulativa ovog pitanja. Zakon će postupno stupati na snagu, s obzirom na složenost materije i neophodnost da se države članice pripreme za njegovu primenu. Svi sistemi veštačke inteligencije koji se danas primenjuju u Evropskoj uniji spadaju u slabu inteligenciju.

3.1. Važenje Zakona

Zakonom su određena lica na koje se odnosi, a u vezi sa preduzimanjem definisanih radnji sa veštačkom inteligencijom na teritoriji važenja. Zakon se primenjuje:

- na ponudioce koji u Uniji stavljaju u promet sisteme veštačke inteligencije ili ih stavljaju u rad, ili modele veštačke inteligencije sa opštim ciljem primene stavljaju u promet, nezavisno od toga da li ovi ponudioci imaju sedište u Uniji ili u nekoj trećoj zemlji;
- na lica koja stavljaju u radi sisteme veštačke inteligencije a imaju sedište u Uniji ili se nalaze u Uniji;
- na ponudioce i lica koja stavljaju u rad sisteme veštačke inteligencije čiji je sedište u nekoj trećoj zemlji, ili se nalaze u trećoj zemlji, onda kada se podaci dobijeni veštačkom inteligencijom primenjuju u Uniji;
- na uvoznike i trgovce sistema veštačke inteligencije;
- na proizvođače koji sisteme veštačke inteligencije zajedno sa njenim produktom stavljaju u promet ili u rad, a pod sopstvenim imenom ili trgovačkim znakom;
- na punomoćnike ponudilaca koji nisu sa sedištem na teritoriji Unije;
- na dotična lica koja se nalaze u Uniji;

Zakon važi samo u područjima koja potpadaju pod pravo Unije i ni u kom slučaju se ne odnosi na nadležnosti država članica u vezi sa nacionalnom bezbednošću, nezavisno od vrste institucija, koje su ovlašćene od strane država članica u pogledu opažanja zadataka i veze sa ovim nadležnostima.

Zakon se ne primenjuje za sisteme veštačke inteligencije kada i ukoliko se isključivo koriste u vojne ciljeve, ciljeve odbrane ili ciljeve nacionalne sigurnosti, a koji su stavljeni u promet, stavljeni u rad i primenjuju se sa ili bez izmena, nezavisno od vrste institucije koja vrši ovu delatnost. Zakon ne važi za sisteme veštačke inteligencije koji nisu u Uniji stavljeni u promet ili u rad, kada se izlazni podaci u Uniji isključivo koriste za vojne svrhe, svrhe odbrane ili nacionalne sigurnosti,

(EU) br. 167/2013, (EU) br. 168/2013, (EU) 2018/858, (EU) 2018/1139 i (EU) 2019/2144, kao i Direktive 2014/90/EU, (EU) 2016/797 i (EU) 2020/1828 (Zakon o veštačkoj inteligenciji) (Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying own harmonised rules on artificial intelligence andamending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)).

nezavisno od vrste ustanove koja ovu delatnost obavlja. Zakon ne važi ni za institucije u trećim zemljama niti za internacionalne organizacije koje prema stavu 1. Zakon su u domenu područja primene ovog zakona, ukoliko ove ustanove ili organizacije upotrebljavaju sisteme veštačke inteligencije u okviru internacionalnog zajedničkog rada ili internacionalnih ugovora u području krivičnog progona i sudske saradnje sa Unijom ili sa jednim ili više članova Unije, i ukoliko je takva treća zemlja ili međunarodna organizacija ponudila primerene garancije u vezi sa zaštitom privatne sfere osnovih prava i sloboda lica.

Ovim Zakonom ne dovodi se u pitanje primena odredaba o odgovornosti pružalaca posrednički usluga u poglavlju 2. Uredbe 2022/2065. Zakon se ne primenjuje na sisteme veštačke inteligencije koji su posebno razvijeni i pušteni u rad isključivo u svrhu naučnog istraživanja i razvoja.

Zakon Unije o zaštiti ličnih podataka, privatnosti i poverljivosti komunikacija primenjuje se na obradu ličnih podataka u vezi sa pravima i obavezama utvrđenim ovim Zakonom[43]. Zakon se ne primenjuje na aktivnosti istraživanja, testiranja i razvoja na sistemima veštačke inteligencije ili AI modelima pre nego što budu stavljeni na tržište ili stavljeni u upotrebu. Takve aktivnosti će se obavljati u skladu sa važećim pravom Unije. Testovi u uslovima stvarnog života ne podležu ovom isključenju. Zakonom se ne utiče na odredbe drugih akata Unije o zaštiti potrošača i bezbednosti proizvoda. Zakon se ne primenjuje na obaveze korisnika koji su fizička lica i koji koriste sisteme u kontekstu čisto, lične i neprofesionalne aktivnosti.

Zakon ne sprečava Uniju ili države članice da održavaju ili uvode zakone i druga akta koja su povoljnija za radnike u pogledu zaštite njihovim prava a u vezi sa korišćenjem sistema veštačke inteligencije, od strane poslodavaca ili da promovišu ili da dozvoljavaju kolektivne ugovore koji su povoljniji za radnike. Zakon se ne primenjuje na sisteme veštačke inteligencije koji su dostupni pod besplatnim licencama, i licencama otvorenog koda osim ako nisu stavljeni na tržište ili upotrebu kao visoko rizični sistemi veštačke inteligencije ili kao sistem obuhvaćen članovima 5. ili 50. Zakona[44].

3.2. Pojam, ciljevi i podela veštačke inteligencije u Zakonu o veštačkoj inteligenciji

Sistem veštačke inteligencije je, u smislu ovog Zakona, sistem zasnovan na mašinama, tako koncipiran da može biti korišćen sa različitim stepenima autonomije i koji posle pokretanja pokazuje sposobnost prilagođavanja koja proizilazi iz primljenih parametara, u eksplicitne ili implicitne svrhe, kao i da pruži izlazne podatke, kao što

[43] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance).

[44] Zakon o vestackoj inteligenciji, čl. 5. i čl. 50.

su predviđanja, sadržaji, preporuke ili odluke koje mogu uticati na fizičko ili virtuelno okruženje. [45]

Cilj navedenog Zakona je poboljšavanje funkcionisanja unutrašnjeg tržišta, promovisanje zaštite za primenu veštačke inteligencije koja je pouzdana i usredsređena na čoveka, uz istovremeno obezbeđivanje visokog nivoa zaštite zdravlja, bezbednosti i osnovnih prava sadržanih u Povelji, uključujući i vladavinu prava, demokratiju i zaštitu životne sredine od štetnih efekata veštačke inteligencije, kao i podrška inovacijama.[46]

Zakonom o veštačkoj inteligenciji se utvrđuje sledeće:[47]

- Harmonizovani propisi za stavljanje u promet i stavljanje u pogon sistema veštačke inteligencije u Evropskoj uniji;
- Zabrana određenih praktika na području veštačke inteligencije;
- Posebni zahtevi u pogledu visoko rizičnih sistema, kao i obaveze aktera u vezi sa takvim sistemima;
- Transparentni propisi za određene sisteme;
- Harmonizovani propisi za stavljanje u promet sistema veštačke inteligencije sa opštom svrhom upotrebe;
- Propisi za opserviranje tržišta, kao i na sprovođenje opserviranja;
- Mere za unapređivanje inovacija sa naročitom pažnjom na preduzeća sa veštačkom inteligencijom uključujući i *startup* kompanije.

Zakon o veštačkoj inteligenciji se zasniva na klasifikaciji rizika koji može nastati njenom primenom. Pri tome je rizik kombinacija verovatnoće nastanka štete i njene težine.[48] Shodno tome, veštačka inteligencija je podeljena na zabranjenu veštačku inteligenciju, visoko rizičnu i minimalno rizičnu.

3.3. Zabranjena veštačka inteligencija

Polazeći od neprihvatljivog rizika koji je stvoren primenom veštačke inteligencije, na području Evropske unije su zabranjene određene primene ili praktike[49]:

- Stavljanje u promet, stavljanje u pogon ili primena sistema veštačke inteligencije kojima se podsvesno, izvan svesti učinioca, ili namerno, manipulativne ili prevarne tehnike upotrebljavaju u cilju ili za uticaj da se postupanje pojedinca ili grupe bitno izmeni tako što se oštećuje njihova sposobnost da donose zasnovane odluke. Na taj način ličnosti su podstaknute da donose odluku koju inače ne bi doneli i to na način koji ovoj osobi, nekoj drugoj ili nekoj grupi nanosi značajnu štetu ili će je tek naneti, odnosno kada deluje na podsvest;

[45] Zakon o veštačkoj inteligenciji, čl. 3. st. 1.
[46] *Isto*, čl. 1. st. 1.
[47] *Isto*, čl. 1. st. 2.
[48] *Isto*, čl. 3. st. 2.
[49] *Isto*, čl. 5. st. 1.

- Stavljanje u promet, stavljanje u pogon ili primena sistema veštačke inteligencije kojima se iskorišćava ranjivost ili potreba za zaštitom nekog fizičkog lica ili određene grupe lica na osnovu starosti, ometenosti ili određene socijalne ili privredne pozicije u cilju ili delovanju da u bitnoj meri promeni odnos ovih lica ili lica koja pripadaju grupi, tako da im se nanosi znatna šteta ili će biti naneta sa velikom verovatnoćom;

- Stavljanje u promet, stavljanje u pogon ili primena sistema veštačke inteligencije za vrednovanje ili stepenovanje fizičkih lica ili grupa lica u određenom vremenskom periodu, na osnovu njihovog socijalnog odnosa ili poznatih, izvedenih, ili programiranih ličnih svojstava, pri čemu socijalno vrednovanje vodi u jednom ili u oba slučaja sledećim rezultatima. To je lošija pozicija ili zapostavljanje određenih fizičkih lica ili grupa lica u socijalnom kontekstu koji nije ni u kakvoj vezi sa okolnostima pod kojima su izvorno dobijeni ili prikupljeni podaci. Pored toga, rezultat može biti lošija pozicija ili zapostavljanje određenih fizičkih lica ili grupa lica, na način koji u pogledu njihovog socijalnog odnosa ili njegovog dosega, nije opravdan, niti je odgovarajući prilikama;

- Stavljanje u promet, stavljanje u pogon nekog sistema veštačke inteligencije u cilju sprovođenja procene rizika u vezi sa fizičkim licem u cilju utvrđivanja rizika da će fizičko lice izvršiti neko krivično delo na osnovu svog profila ili ocene njegovih personalnih obeležja ili osobina. Ova zabrana ne važi za sisteme veštačke inteligencije koji se koriste u cilju podrške ljudskoj proceni o umešanosti osobe u kriminalnu aktivnost koja je već zasnovana na objektivnim i proverljivim činjenicama, koje su u direktnoj vezi sa nekom kriminalnom delatnošću;

- Stavljanje u promet, stavljanje u pogon za tu specifičnu svrhu ili korišćenje za utvrđivanje emocija fizičkog lica na radnom mestu i u obrazovnim institucijama, osim kada je upotreba sistema namenjena za stavljanje na tržište iz medicinskih ili bezbednosnih razloga;

- Stavljanje na tržište, stavljanje u pogon u tu posebnu svrhu ili korišćenje sistema biometrijske kategorizacije, koji individualno kontrolišu fizička lica na osnovu njihovih biometrijskih podataka, kako bi se zaključilo o njihovoj rasi, političkom mišljenju, članstvu u sindikatima, verskim ili filozofskim uverenjima, seksualnom životu ili orjentaciji. Ova zabrana se ne odnosi na obeležavanje ili filtriranje stečenih biometrijskih podataka, na pravno dozvoljen način, kao što su slike na temelju biometrijskih podataka ili kategorisanja biometrijskih podataka na području krivičnog gonjenja;

- Korišćenje daljinskih biometrijskih sistema za identifikaciju u realnom vremenu u javno dostupnim prostorijama u svrhe sprovođenja zakona, osim kada je i u meri u kojoj je to neophodno za postizanje određenih ciljeva. To su ciljane potrage za konkretnim žrtvama otmice, trgovine ljudima ili seksualne eksploatacije kao i traženje nestalih lica. Pored toga, reč je i o sprečavanju konkretne, značajne i neposredne opasnosti po život ili fizički integritet lica ili stvarne ili predvidljive pretnje terorističkim napadom. Takođe, reč je o potrebi lociranja ili identifikovanja lica za koje se sumnja da je počinilo krivično delo, a u svrhe vođenja krivičnog

postupka ili postupka za izvršenje kazne za dela navedena u odgovarajućem delu Aneksu II, a koja su kažnjiva u državi članici, prema njenom zakonodavstvu, kaznom zatvora ili nalogom za pritvor od najmanje 4 godine.

3.4. Sistemi visokog rizika

Pored sistema veštačke inteligencije čija je upotreba zabranjena, Zakonom o veštačkoj inteligenciji su predviđena posebna pravila za primenu koja se smatra visoko rizičnom. To su sistemi koji su posebno rizični za zdravlje i sigurnost ili za osnovna prava fizičkih lica. Ovi sistemi su podeljeni u dve osnovne kategorije.[50] To su:

- Sistemi koji će biti primenjeni u proizvodima koji su obuhvaćeni propisima Evropske unije u vezi sa sigurnošću proizvoda. Ovde spadaju igračke, automobili, medicinski uređaji i liftovi, kao i oni koji se odnose na vazdušni saobraćaj.
- Sistemi koji spadaju u specifična područja i koja moraju biti registrovana u nekoj od bazi podataka u Evropskoj uniji. Konkretno, reč je o:
  - Sigurnosne komponente u kritičnim infrastrukturama, npr. saobraćaju, čiji iskakanje može ugroziti život i zdravlje građana;
  - Rešenja veštačke inteligencije koja se primenjuju u obrazovnim institucijama i kojima se može definisati pristup obrazovanju i tok karijere nekog lica, npr. vrednovanje ispita;
  - Sigurnosne komponente proizvoda koje se baziraju na veštačkoj inteligenciji, npr. primena u hirurgiji potpomognuto robotima;
  - Alati za zapošljavanje, menadžment u vezi sa zaposlenima i pristup samostalnosti, npr. kompjuterski softver za sortiranje u vezi sa zapošljavanjem;
  - Određeni slučajevi primene veštačke inteligencije koji se koriste za omogućavanje pristupa bitnim privatnim ili javnim uslugama, npr. procena kreditne sposobnosti kojima se uskraćuje mogućnost građanima da dobiju kredit;
  - Sistemi za biometrijsko daljinsko identifikovanje, prepoznavanje emocija i biometrijsko kategorizovanje, npr. sistemi za povratno identifikovanje neke građe;
  - Slučajevi primene u krivičnom postupku koji mogu zadreti u osnovna ljudska prava, npr. ocena pouzdanosti dokaznih sredstava;
  - Slučajevi primene u menadžmentu migracija, azila i graničnih kontrola, npr. automatizovanje ispitivanja zahteva za vizu;
  - Rešenja veštačke inteligencije u oblasti demokratskih procesa i organizovanja pravosuđa, npr. rešenja veštačke inteligencije za pripremu sudskih presuda.

Sistemi visokog rizika potpadaju strogim obavezama pre nego što su stavljeni u promet. [51] To se odnosi na: primerenu ocenu rizika i sisteme za smanjenje rizika;

[50] Zakon o veštačkoj inteligenciji, čl. 6. st. 1.

[51] KI-Gesetz: erste Regulierung der künstlichen Intelligenz, *https://www.europarl.europa.eu /topics/de/article/20230601STO93804/ki-gesetz-erste-regulierung-der kunstlichen-intelligenz*, 02.04.2025.

visoki kvalitet memorisanih podataka u cilju minimalizacije rizika diskriminirajućih rezultata; protokolisanje aktivnosti u cilju garantovanja povratnog praćenja rezultata; obimna dokumentacija koja sadrži neophodne informacije o sistemu i njegovoj svrsi, tako da nadležni organi mogu da ocene konformnost sistema; jasne i primerene informacije o licu preduzetnika; primerene mere nadzora koje preduzima čovek; visoka izdržljivost, sigurnost u pogledu sajber bezbednosti i tačnost.

3.5. Minimalni rizik i zahtevi u pogledu transparentnosti

U zakonu nisu sadržane odredbe za veštačku inteligenciju kojom se izaziva minimalni ili nikakav rizik. Većina sistema veštačke inteligencije koji se danas primenjuju u Evropskoj uniji spada u ovu kategoriju. Npr. primena veštačke inteligencije za video igre i neželjene poruke poslate elektronskom poštom.

Modeli kao što je *ChatGPT* nisu svrstani u kategoriju visokog rizika, ali moraju da ispune zahteve u vezi transparentnosti i autorskog prava Evropske unije.[52] Ti zahtevi su: objavljivanje da je sadržaj generisan veštačkom inteligencijom, oblikovanje modela na način kojim se sprečava da budu dobijeni ilegalni sadržaji; objavljivanje sažetaka podataka koji su zaštićeni autorskim pravom, a upotrebljeni su za trening sistema. Sistemi sa opštim ciljem primene i značajnim delovanjem koji bi mogao da predstavlja sistemski rizik, kao što su *ChatGPT 4*, moraju temeljno da budu vrednovani i sve teške posledice trebalo bi prijaviti Komisiji.

Sadržaji koji su dobijeni primenom veštačke inteligencije ili njome promenjeni, kao što su fotografije, vizuelni ili audio podaci moraju jasno biti označeni kao generisani veštačkom inteligencijom tako da korisnici znaju da dolaze u kontakt sa takvim sadržajima.

3.6. Stupanje na snagu i primena Zakona

Zakon stupa na snagu 1.8.2024. godine, i 2.8.2026. godine biće primenjen u punom obimu. Izuzeci su zabrana sistema sa neprihvatljivim rizikom stupa na snagu 6 meseci posle stupanja na snagu zakona; propisi za sisteme sa opštom primenom, koji moraju odgovarati zahtevima transparentnosti, stupaju na snagu 12 meseci posle stupanja na snagu zakona; sistemi sa visokim rizikom, s obzirom na mere koje moraju biti preuzete, stupaju na snagu 36 meseci posle stupanja na snagu zakona.

U februaru 2024. godine osnovan je Evropski ured za veštačku inteligenciju. On nadgleda sprovođenje i implementaciju zakona u državama članicama. Evropski parlament je osnovao specijalnu radnu grupu koja nadgleda sprovođenje i poštovanje zakona. Na taj način treba da bude obezbeđeno da se propisima postojano može podsticati razvoj digitalnog sektora u Evropi. Osnovana su i 3 savetodavna tela. To su

[52] KI-Gesetz, *https://digital-strategy.ec.europa.eu/de/policies/regulatory-framework-ai*, 02.04. 2025; KünstlicheIntelligenz, *https://digital-strategy.ec.europa.eu/de/policies/artificial-intelligence*, 02.04.2025.

Evropski odbor za veštačku inteligenciju koji se sastoji od predstavnika zemalja članica, Naučni gremijum koji čine nezavisni eksperti na području veštačke inteligenciji i Savet koji zastupa interese kako komercijalnih, tako i nekomercijalnih učesnika[53].

4. Zaključak

Veštačka inteligencija više nije samo sredstvo kojim se čovek koristi za postizanje definisanih ciljeva. Napredniji oblici, već sada, deluju u nekoj meri autonomno, tako da se izlazni rezultat često ne može predvideti. Međutim, činjenica da istraživači nisu u mogućnosti da objasne procese koji se odvijaju u veštačkoj inteligenciji, već po sebi nosi opasnost da se oni ne mogu kontrolisati. Još viši nivoi veštačke inteligencije koji će delovati sasvim autonomno, podižu opasnost neočekivanih rezultata na čovečanstvo. Ispitivanje funkcionisanja veštačke inteligencije odvija se paralelno i u vezi sa istraživanjima načina funkcionisanja ljudskog mozga. Napredak u ovoj oblasti ima bitan uticaj na razumevanje same veštačke inteligencije i omogućava njeno dalje napredovanje. Svakako, veštačka inteligencija samo u nekim aspektima podražava čovekov mozak. Zato je vrlo teško očekivati stvaranje veštačke inteligencije koja bi bila samostalna, posedovala empatiju ili mogla da razlikuje dobro od zla.

Veštačka inteligencija je izuzetno značajna dopunska tehnologija kojom se poboljšavaju čovekove sposobnosti. Ljudi mogu da postupno umanje svoju sposobnost za samostalnim razmišljanjem i kreacijama u slučaju čestog oslanjanja na veštačku inteligenciju u rešavanju mnogih zadataka. Iako je reč o značajnom napretku i usavršavanju, ne mogu se izbeći mnoge opasnosti koje se izazivaju njenom primenom. Neke od značajnih prikazane su u ovom radu. Na primer,. veštačka inteligencija može da pogrešno proceni neku situaciju i da deluje na način na koji se izazivaju teške posledice. Kada je reč o autonomnim autima, posledica može biti gubitak života. Ili raketni odbrambeni sistemi, mogu u slučaju greške izazvati razorne posledice po čovečanstvo. Sve to znači da bi takvi sistemi morali da poseduju odgovarajuće zaštitne mehanizme kojima se eliminiše ili smanjuje pogrešna primena.

Aktuelna svetska situacija pokazuje i direktnu težnju da se veštačka inteligencija, suprotno izvornoj nameri, koristi u ratovima sa posledicama koje se u ovom trenutku ne mogu proceniti. Pored toga, ulaganje u razvoj i primenu veštačke inteligencije su takva da ih mogu podneti samo najrazvijenije zemlje. U tom smislu, postoji duboka opasnost da manje razvijene zemlje koje nisu u stanju da ulažu sredstava, dolaze u vid novog oblika kolonijalizma koji bi se mogao nazvati kolonijalizam na osnovu veštačke inteligencije. Tačno je da su se i dosadašnji oblici kolonijalizma zasnivali na tehnološkoj superiornosti svetskih sila, no ovde je reč o mnogo direktnijem, snažnijem delovanju koje ima za posledicu, kako smo istakli, specifično tehnološko ropstvo.

[53] Zakon o vestackoj inteligeniji, čl. 113.

Neophodno je istaći da su svi modeli veštačke inteligencije danas vezani za kvalitet i kvantitet podataka kojima se treniraju. U slučaju da podaci nisu dovoljni ili kada je reč o pogrešnim i nepotpunim podacima, veštačka inteligencija ne može da iskaže pouzdane rezultate. Zato je jedno od suštinskih pitanja potpuniji i što je moguće sveobuhvatniji trening, kao i jasno i odgovorno definisanje ciljeva i usavršavanje snažnijih mehanizama, naročito u oblastima koje se smatraju visokorizičnim.

Pravna regulativa veštačke inteligencije je tek na samom početku. U Evropskoj uniji su naglašeni napori da se neka pitanja pravno urede, pa je 1donet Zakon o veštačkoj inteligenciji, koji je rezultat višegodišnjeg istraživanja ovog fenomena. Mnoga od značajnih pitanja nisu obuhvaćena ovim Zakonom. Međutim, najznačajnije je to da je principijelno rešen odnos društva prema veštačkoj inteligenciji. On je zasnovan na visini rizika, tj. na štetnim posledicama koje se mogu izazvati veštačkom inteligencijom, pa je ona podeljena na onu koja je zabranjena, onu koju izaziva visok rizik i onu koju ne izaziva rizik.

S veštačkom inteligencijom se povezuje i veliki broj drugih pitanja koja nisu predmet regulative Zakona. Npr. bitno je pitanje oblika pravne zaštite veštačke inteligencije kao takve, patentiranje kompjuterskih programa koji se nalaze u njenoj osnovi, način na koji će se štititi proizvodi dobijeni veštačkom inteligencijom. Adekvatnim pravnim rešenjima mora biti podržana veštačka inteligencija, ali uzeta u obzir i potreba društva i pojedinaca za slobodnim razvojem konkurencije.

Milica Šutova, Ph.D., Associate Professor
Faculty of Law, Goce Delčev University in Štip
North Macedonia

Ksenija Paunović, Ph.D., Scientist Associate
University of Kragujevac

SOCIALRISKSOFARTIFICIALINTELLIGENCEAPPLICATION ANDITSREGULATIONINEUROPEANUNIONLAW

Summary

Artificial intelligence is being applied in almost all areas of life. Legal regulation is unable to keep pace with its rapid and multifaceted development. The experience to date with the use of existing weak artificial intelligence has led to noticeable and serious societal risks. For this reason, the European Union has adopted the Artificial

Intelligence Act, which regulates the fundamental relationship of society towards this phenomenon. The Act is based on the level of risk associated with its application; thus, certain forms are considered prohibited, while high-risk artificial intelligence is permitted, subject to stringent requirements for its acceptance. At the same time, the subject matter of this Act is not to define rights, obligations, or legal protection relating to artificial intelligence itself as a subject of law.

Key words: *weak artificial intelligence, strong artificial intelligence, prohibited artificial intelligence, high-risk systems, transparency*

Literatura

Akula, A.R., Wang, K., Liu, C., Saba-Sadiya, S., Lu. H., Todorovic, S., Chai, J., Zhu, S.C., CX-*ToM: counterfactual explanations with theory-of-mind for enhancing human trust in image recognition models*, i-Science 25(1) 2022.

Alpaydin, E., *Introduction to Machine Learning (4th Edition)*, MIT Press, 2020, dostupno na: https://mitpress.mit.edu/books/introduction-machine-learning-fourth-edition

Acemoglu, D., Restrepo, P., *Artificial Intelligence, Automation, and Work*, NBER Working Paper No. 24196, 2018, dostupno na: https://www.nber.org/papers/w24196

Armstrong, S., Sandberg, A., Bostrom, N., T*hinking inside the box: Controlling and using an oracle AI*, Minds and Machines, 22(4) 2012.

Ji, Z. *et. al., Survey of Hallucination in Natural Language Generation*, ACM Computing Surveys (CSUR), 55(12) 2022.

Kariluoto, A., Kultanen, J., Soininen, J., Pärnänen, A., Abrahamsson, P., *Quality of Data in Machine Learning*, 2021, dostupno na: DOI: 10.1109/QRS-C55045.2021.00040

Kar, K., Kornblith, S., Fedorenko, E., *Interpretability of artificial neural network models in artificial intelligence vs. Neuroscience*, dostupno na: DOI:10.48550/arXiv.2206.03951

Li, L., Allamanis, M., Brockschmidt, M., Gaunt, A., *Competition-Level Programming Challenge with AlphaCode*, Nature, 607/2022.

Mavrogiorgos, K., Kiourtis, A., Mavrogiorgou, A., Menychtas, A., Kyriazis, D., *Bias in Machine Learning: A Literature Review, Department of Digital Systems,* dostupno na: https://doi.org/10.3390/app14198860

Nguyen, T. T. *et. al., Deep Learning for Deepfakes Creation and Detection: A Survey*, dostupno na: https://doi.org/10.48550/arXiv.1909.11573

Momand, Z., Chan, H. J., Mongkolnam, P., *Immersive Technologies in Virtual Companions:A Systematic Literature Review*, dostupno na: https://www.researchgate.net/publication/373686231_Immersive_Technologies_in_Virtual_Companions_A_Syematic_Literature_Review

Santoro, A., Lampinen, A., Mathewson, K., Lillicrap, T., Raposo, D., *Symbolic Behaviour in Artificial Intelligence,* 2022, dostupno na: https://doi.org/10.48550/arXiv.2102.03406

Skoudis, E., Zeltser, L., *Malware: Fighting Malicious Code*, Prentice Hall PTR, New Jersey, 2004.

Sutton, R. S., Barto, A. G., *Reinforcement Learning: An Introduction (2nd edition)*, MIT Press, 2018, dostupno na: http://incompleteideas.net/book/the-book-2nd.html

Frey, C. B., Osborne, M. A., *The future of employment: How susceptible are jobs to computerisation?* Technological Forecasting and Social Change, 114/2017.

Fritsch, L., Jaber, A., Yazidi, A., *An Overview of Artificial Intelligence Used in Malware*, In: Zouganeli, E., Yazidi, A., Mello, G., Lind, P. (eds) Nordic Artificial Intelligence Research and Development. NAIS 2022. Communications in Computer and Information Science, vol 1650. Springer, Cham, dostupno na: https://doi.org/10.1007/978-3-031-17030-0_4

Chen, M., Mao, S., Liu, Y., *Big Data: A Survey*, Mobile Networks and Applications, 19(2) 2014.ArtenvonKI: Definition, Abgrenzung&Anwendung, Otrissoftware, *https://www.otris.de/wiki/ arten-von-ki/*, 10.4.2025.

GefahrendurchKI: EinBlickaufdieRisikenderkünstlichenIntelligenz, https://www.moin.ai/chatbot - lexikon/gefahren-durch-ki, 07.04.2025.;

DieGefahrenundRisikenvonKünstlicherIntelligenz (KI), https://www.ki-company.ai, 07.04.2025.

KI-SystemeKönnenLügenUndBetrügen, https://www.it-daily.net/shortnews/ki-systeme-koennen-luegen-und-betruegen, 07.04.2025.

KI-Gesetz: ersteRegulierungderkünstlichenIntelligenz, https://www.europarl.europa.eu/topics/de/article/20230601STO93804/ki-gesetz-erste-regulierung-der-kunstlichen-intelligenz, 02.04.2025.

KI-Gesetz, https://digital-strategy.ec.europa.eu/de/policies/regulatory-framework-ai, 02.04.2025.

KünstlicheIntelligenz, https://digital-strategy.ec.europa.eu/de/policies/artificial-intelligence, 02.04.2025.

KünstlicheIntelligenzundDigitalisierung: WelcheChancenundRisikengibtes?, https://ambient.digital/wissen/blog/kuenstliche-intelligenz-chancen-risiken/, 07.04.2025.

Regulativa (EU) 2024/1689 Evropskog parlamenta i Savetaod 13. juna 2024. godine, kojom se utvrđuju usklađena pravila o veštačkoj inteligenciji i menja Regulativa (EZ) br. 300/2008, (EU) br. 167/2013, (EU) br. 168/2013, (EU) 2018/858, (EU) 2018/1139 i (EU) 2019/2144, kao i Direktive 2014/90/EU, (EU) 2016/797 i (EU) 2020/1828 (Zakon o veštačkoj inteligenciji) (Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and a mending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)).

Chancen, GrenzenundRisikenderkünstlichenIntelligenz – vonihrselbsterklärt, https://www.argus datainsights. ch/de/blog/chancen-grenzen-und-risiken-der-kuenstlichen-intelligenz-von-ihr-selbst-erklaert, 07.04.2025.